can analogs of such proteins. In addition, completely novel polypeptides can also be synthesized, as can proteins incorporating non-naturally occurring amino acids.

In a protein, the peptide bonds between adjacent amino acid residues are resonance hybrids of two different electron isomeric structures, wherein a bond between a carbonyl carbon (the carbon atom of the carboxylic acid group of one amino acid after its incorporation into a protein) and a nitrogen atom of the amino group of the α-carbon of the next amino acid places the carbonyl carbon approximately 1.33 Å away from the nitrogen atom of the next amino acid, a distance about midway between the distances that would be expected for a double bond (about 1.25 Å) and a single bond (about 1.45 Å). This partial double bond character prevents free rotation of the carbonyl carbon and amino nitrogen about the covalent bond therebetween under physiological conditions. As a result, the atoms bonded to the carbonyl carbon and amino nitrogen reside in the same plane, and provide discrete regions of structural rigidity, and hence conformational predictability, in proteins.

Beyond the peptide bond, each amino acid residue contributes two additional single covalent bonds to the polypeptide chain. While the peptide bond limits rotational freedom of the carbonyl carbon and the amino nitrogen of adjacent amino acids, the single bonds of each residue (between the α-carbon and carbonyl carbon (the phi (φ) bond) and between the α-carbon and amino nitrogen (the psi (ψ) bond) of each amino acid residue), have greater rotational freedom. For example, the rotational angles for φ and ψ bonds for certain common regular secondary structures are listed in the following table:

Structure	Approximate Bond Angle		Residues per turn	Helix pitch (Å)[a]
	φ	ψ		
Right-handed α-helix (3.6_{13} - helix)	-57	-47	3.6	5.4
3_{10} - helix	+49	-26	3.0	6.0
Parallel β-strand	-119	+113	2.0	6.4
Antiparallel β-strand	-139	+135	2.0	6.8

[a] "Helix pitch" refers to the distance between repeating turns on a line drawn parallel to the helix axis. Bond angles associated with other secondary structures are known in the art, or can be determined experimentally using standard techniques.

Similarly, the single bond between a α-carbon and its attached R-group provides limited rotational freedom. Collectively, such structural flexibility enables a number of possible conformations to be assumed at a given region within a polypeptide. As discussed in greater detail below, the particular conformation actually assumed depends on thermodynamic considerations, with the lowest energy conformation being preferred.

In addition to primary structure, proteins also have secondary, tertiary, and, in multi-subunit proteins, quaternary structure. "Secondary structure" refers to local conformation of the polypeptide chain, with reference to the covalently linked atoms of the peptide bonds and α-carbon linkages that string the amino acid residues of the protein together. Side chain groups are not typically included in such descriptions. Representative examples of secondary structures include α helices, parallel and anti-

parallel β structures, and structural motifs such as helix-turn-helix, β-α-β, the leucine zipper, the zinc finger, the β-barrel, and the immunoglobulin fold. Movement of such domains relative to each other often relates to biological function and, in proteins having more than one function, different binding or effector sites can be located in different domains.

"Tertiary structure" concerns the overall three-dimensional structure of a protein, including the spatial relationships of amino acid residue side chains and the geometric relationship of different regions of the protein. "Quaternary structure" relates to the structure and non-covalent association of different polypeptide subunits in a multisubunit protein.

A "functional site" refers to any site in a protein that has a function. Representative examples include active sites (*i.e.*, those sites in catalytic proteins where catalysis occurs), protein-protein interaction sites, sites for chemical modification (*e.g.*, glycosylation and phosphorylation sites), and ligand binding sites. Ligand binding sites include, but are not limited to, metal binding sites, co-factor binding sites, antigen binding sites, substrate channels and tunnels, and substrate binding sites. In an enzyme, a ligand binding site that is a substrate binding site may also be an active site.

A "pseudoatom" refers to a position in three dimensional space (represented typically by an x, y, and z coordinate set) that represents the average (or weighted average) position of two or more atoms in a protein or amino acid. Representative examples of a pseudoatom include an amino acid side chain center of mass and the center of mass (or, alternatively, the average position) of an α-carbon atom and the carboxyl atom bonded thereto. Hypothetical covalent bonds between pseudoatoms, or between a pseudoatom and another atom, are referred to herein as "virtual covalent bonds."

A "geometric constraint" or "tertiary constraint" refers to a spatial parameter with respect to an atom or group of atoms (*e.g.*, an amino acid, the R-group of an amino acid, the center of mass of an R-group of an amino acid, a pseudoatom, *etc.*).